PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 4 : C07H 21/02 — A1

(11) International Publication Number: WO 86/ 06726

(43) International Publication Date: 20 November 1986 (20.11.86)

(21) International Application Number: PCT/US86/01052

(22) International Filing Date: 15 May 1986 (15.05.86)

(31) Priority Application Number: 734,323

(32) Priority Date: 15 May 1985 (15.05.85)

(33) Priority Country: US

(71) Applicant: INTEGRATED GENETICS, INC. [US/US]; 51 New York Avenue, Framingham, MA 01701 (US).

(72) Inventor: MOCK, Graham, A. ; 14 Mount Vickery Road, Southboro, MA 01772 (US).

(74) Agent: CLARK, Paul, T.; Fish & Richardson, One Financial Center, Boston, MA 02111 (US).

(81) Designated States: AT (European patent), BE (European patent), CH (European patent), DE (European patent), FR (European patent), GB (European patent), IT (European patent), JP, LU (European patent), NL (European patent), SE (European patent).

Published
With international search report.

(54) Title: CYTIDINE ANALOGS

(57) Abstract

An oligonucleotide containing at least one cytidine analog labeled at C4 with a reporter group.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GA**	Gabon	**MR**	Mauritania
AU	Australia	**GB**	United Kingdom	**MW**	Malawi
BB	Barbados	**HU**	Hungary	**NL**	Netherlands
BE	Belgium	**IT**	Italy	**NO**	Norway
BG	Bulgaria	**JP**	Japan	**RO**	Romania
BR	Brazil	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SN**	Senegal
CH	Switzerland	**LI**	Liechtenstein	**SU**	Soviet Union
CM	Cameroon	**LK**	Sri Lanka	**TD**	Chad
DE	Germany, Federal Republic of	**LU**	Luxembourg	**TG**	Togo
DK	Denmark	**MC**	Monaco	**US**	United States of America
FI	Finland	**MG**	Madagascar		
FR	France	**ML**	Mali		

CYTIDINE ANALOGS

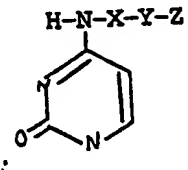
Background of the Invention

This invention relates to nucleotide analogs capable of being attached to oligonucleotides and labeled for use as nucleic acid hybridizati n probes.

Such probes are used to detect DNA and, in some instances, RNA, characteristic of particular cells or microorganisms. Both isotopic and non-isotopic labeling methods have been used.

Summary of the Invention

In general, the invention features a cytidine analog in which the cytosine base portion thereof has the formula



wherein Y is derived from a group capable of reacting with the N-hydroxy-succinimide ester of biotin amino caproic acid, Z is a protecting group capable of preventing Y from reacting with a phosphite, and X is a linking group connecting Y and N.

Preferably, Y is NH, Z is $\overset{\overset{O}{\|}}{C}-CF_3$, and X is $(CH_2)_3$.

As will be explained in more detail below, the cytidine analogs of the invention are attached to synthetic oligonucleotide probes, and the protective group Z removed to expose the reactive group Y, which is spaced from the cytidine analog by spacer X, to prevent steric hindrance. A reporter group is then attached to reactive group Y (which may change slightly in structure in the process of deprotection; e.g., NH becomes NH_2), thus labeling the probe.

The invention provides rapid, simple labeling which can be precisely controlled in degree and position of label.

Other features and advantages of the invention will be apparent from the following description of the preferred embodiments thereof, and from the claims.

Description of the Preferred Embodiments

The drawings will first be described.

Drawings

Fig. 1 is the structure of a preferred labeled oligonucleotide of the invention.

Figs. 2-5 are the structures of the intermediates in the synthesis of said labeled oligonucleotide.

Structure

The labeled oligonucleotides of the invention have the structure given in the Summary of the Invention, above. The substituents are now described in more detail.

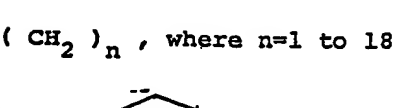
Reactive Group Y

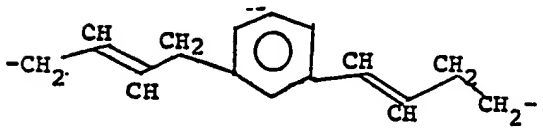
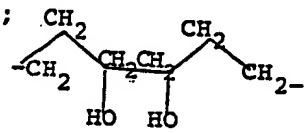
Reactive group Y is derived from a group capable of reacting with the N-hydroxy-succinimide ester of biotin amino caproic acid. This does not mean that the reporter group need be biotin or a derivative thereof, but is simply a convenient way of defining the reactivity of Y necessary for its being labeled with any desired reporter group.

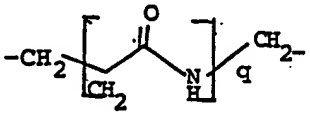
The presently most preferred reactive group Y is an amino group or derivative thereof. Other suitable reactive groups are S- and  .

Linking Group X

Linking group X does not participate directly in the derivatization reaction but rather serves the function of preventing steric hindrance between the amino group of cytidine and reactive group Y. Consequently, it is the size of X, rather than its chemical composition, that is important. Generally, X therefore can include any of a wide range of chemical substituents, provided they are not more reactive than group Y. Examples of suitable linking groups X are

$(CH_2)_n$, where n=1 to 18 ;  where p= 1 to 6 ;

 ;  ; and

 where q= 1 to 9.

Protecting Group Z

Protecting group Z is any group which bonds to reactive group Y to prevent Y from reacting with a phosphite reagent. The protecting group Z must also be capable of being removed after the cytidine analog has been attached to the oligonucleotide, under conditions which do not substantially impair the hybridizing ability of the oligonucleotide or the reactivity of Y.

Where the reactive group Y is an amino group, suitable protecting groups are amides, e.g., N-Formyl, N-Acetyl, N-Chloroacetyl, N-Trichloroacetyl, N-Trifluoroacetyl, N-o-Nitrophenylacetyl, N-o-Nitrophenoxyacetyl, N-Acetoacetyl, N-3-Phenylpropionyl, N-3-(p-Hydroxyphenyl)propionyl,

N-2-Methyl-2-(-nitrophenoxy)propionyl, N-2-Methyl-2-(o-phenylazophen xy)propionyl, N-4-Chlorobutyryl, N-o-Nitrocinnamoyl, N-Picolinoyl, N-(N'-Acetylmethionyl), N-Benzoyl, N-Phthaloyl, and N-Dithiasuccinoyl.

Other suitable amino protecting groups are carbamates, e.g, Methyl Carbamate, 9-Fluorenylmethyl Carbamate, 2,2,2-Trichloroethyl Carbamate, 2-Trimethylsilylethyl Carbamate, 1,1-Dimethylpropynyl Carbamate, 1-Methyl-1-phenylethyl Carbamate, 1-Methyl-1-(4-biphenylyl)ethyl Carbamate, 1-1-Dimethyl-2-haloethyl Carbamate, 1,1-Dimethyl-2-cyanoethyl Carbamate, t-Butyl Carbamate, Cyclobutyl Carbamate, 1-Methylcyclobutyl Carbamate, 1-Adamantyl Carbamate, Vinyl Carbamate, Allyl Carbamate, Cinnamyl Carbamate, 8-Quinolyl Carbamate, N-Hydroxypiperidinyl Carbamate, 4,5-Diphenyl-3-oxazolin-2-one, Benzyl Carbamate, p-Nitrobenzyl Carbamate, 3,4-Dimethoxy-6-nitrobenzyl Carbamate, 2,4-Dichlorobenzyl Carbamate, 5-Benzisoxazolylmethyl Carbamate, 9-Anthrylmethyl Carbamate, Diphenylmethyl Carbamate, Isonicotinyl Carbamate, S-Benzyl Carbamate, and N-(N'-Phenylaminothiocarbonyl) Derivatives.

Other suitable amino protecting groups include N-Allyl, N-Phenacyl, N-3-Acetoxypropyl, Quaternary Ammonium Salts, N-Methoxymethyl, N-Benzyloxymethyl, N-Pivaloyloxymethyl, N-Tetrahydropyranyl, N-2,4-Dinitrophenyl, N-Benzyl, N-o-Nitrobenzyl, N-Di(p-methoxyphenyl)methyl, N-Triphenylmethyl, N-(p-Methoxyphenyl)diphenylmethyl, N-Diphenyl-4-pyridylmethyl, N-2-Picolyl N'-Oxide, N,N'-Isopropylidene, N-Benzylidene, N-p-Nitrobenzylidene, N-Salicylidene, N-(5,5-Dimethyl-

3- x -1-cyclohexenyl), N-Nitro, N-Oxide, N-Diphenylphosphinyl, N-Dimethylthioph sphinyl, N-Benzenesulfenyl, N-o-Nitrobenzenesulfenyl, N-2,4,6-Trimethylbenzenesulfonyl, N-Toluenesulfonyl, N-Benzylsulfonyl, N-Trifluoromethylsulfonyl, and N-Phenacylsulfonyl.

Where reactive group Y is a thiol group, suitable protecting groups include S-Benzyl Thioether, S-p-Methoxybenzyl Thioether, S-p-Nitrobenzyl Thioether, S-4-Picolyl Thioether, S-2-Picolyl N-Oxide Thioether, S-9-Anthrylmethyl Thioether, S-Diphenylmethyl Thioester, S-Di(p-methoxyphenyl)methyl Thioether, S-Triphenylmethyl Thioether, S-2,4-Dinitrophenyl Thioester, S-t-Butyl Thioether, S-Isobutoxymethyl Hemithioacetal, S-2-Tetrahydropyranyl Hemithioacetal, S-Acetamidomethyl Aminothioacetal, S-Cyanomethyl Thioether, S-2-Nitro-1-Phenylethyl Thioether, S-2,2-Bis(carboethoxy)ethyl Thioether, S-Benzoyl Derivative, S-(N-Ethylcarbamate), and S-Ethyl Disulfide.

Where the reactive group Y is a carboxyl group, suitable protecting groups include Methyl Ester, Methoxymethyl Ester, Methylthiomethyl Ester, Tetrahydropyranyl Ester, Benzyloxymethyl Ester, Phenacyl Ester, N-Phthalimidomethyl Ester, 2,2,2-Trichloroethyl Ester, 2-Haloethyl Ester, 2-(p-Toluenesulfonyl)ethyl Ester, t-Butyl Ester, Cinnamyl Ester, Benzyl Ester, Triphenylmethyl Ester, Bis(o-nitrophenyl)methyl Ester, 9-Anthrylmethyl Ester, 2-(9,10-Dioxo)anthrylmethyl Ester, Piperonyl Ester, Trimethylsilyl Ester, t-Butyldimethylsilyl Ester, S-t-Butyl Ester, 2-Alkyl-1,3-oxazolines, N,N-Dimethylamide, N-7-Nitroindoylamide, Hydrazides, N-Phenylhydrazide, and N,N-Diisopropylhydrazide.

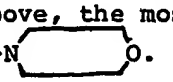
Sugar Moiety

For convenience and economy, it is preferred that the sugar moiety to which the cytosine base is bonded have the conventional 5-membered heterocyclic configuration, with oxygen and four carbon atoms constituting the ring. Strict adherence to this configuration is not important to the labeling or hybridization reactions, however, and various other substituents, e.g., S, NH, or CH_2, can be substituted for the ring oxygen if desired.

Similarly, the group bonded to C_5 of the sugar moiety is most conveniently a hydroxyl group in which the hydrogen is replaced by a protecting group which is capable of preventing the reaction of the hydroxyl with a phosphite reagent; this protecting group can be any conventional hydroxyl protecting group, e.g., 4,4'-dimethoxytrityl (DMT), ethers such as methyl ether, esters such as acetate esters, and cyclic acetals and ketals such as ethylidine.

If desired, the group attached to C_5 of the sugar moiety can be, rather than a protected hydroxyl group, a conventionally protected S group.

Linkage to Oligonucleotide

The cytidine analog is attached to the oligonucleotide by means of a linking group substituted for the hydrogen in the hydroxyl group bonded to C_3 of the sugar moiety. As mentioned above, the most preferred linking group is $CH_3O\text{-}P\text{-}N$  . Any other suitable phosphorus-containing group can be used, however, e.g., other phosphite groups, or phosphate groups, including groups containing more than one phosphorus atom.

Reporter Group

Following the attachment of the cytidine analog to the oligonucle tide and the deprotecti n of reactive group Y, Y is reacted with any suitable reporter group. The reporter group can be directly detectable, as in the case of fluorophores such as fluorescein, or it can be a group which confers binding specificity on the oligonucleotide to enable detection by a subsequent compound or series of compounds. One well-known example of the latter is the system using biotin, avidin, an enzyme, and a substrate for the enzyme which produces a detectable change when acted on by the enzyme.

Synthesis

The synthesis of the labeled oligonucleotides of the invention is carried out generally as follows.

The first step is to synthesize a cytidine analog having the formula given in the Summary of the Invention, above. The analog has, attached to the amino group at C4, a linking group X, a reactive group, or derivative thereof, Y, and a protective group Z. To construct the analog, X and Y are attached to the amino group of C4 of cytidine, and the reactive group Y is then protected with protective group Z.

The next step is to attach the cytidine analog to the 5' end of an oligonucleotide to be used as a hybridization probe. Following this step, additional nucleotides can be attached to the cytidine analog, including additional cytidine analogs, to give as many labeled positions in the oligonucleotide as desired, at precisely predetermined positions. This is conveniently done using commercially available DNA synthesizers, int which the cytidine analog has been loaded.

Finally, the cytidine analog is deprotected and a label attached at Y to yield a labeled oligonucleotide.

Referring now to Fig. 1, there is shown a labeled oligonucleotide in which X is $(CH_2)_3$, Y is NH, and the detectable label is biotin, a well-known reporter compound whose affinity for avidin has previously rendered it useful in labeling DNA probes.

Figs. 2-5 show the synthetic steps in the construction of the labeled oligonucleotide of Fig. 1. Referring to Fig. 2, the addition of the linker $(CH_2)_3$ and reactive group NH_2 is accomplished in a single step. This is followed by the addition to the NH_2 of the protective group $COCF_3$. The next step is protection of the C5 hydroxyl with the protective group DMT. The remaining hydroxyl (C3) is then reacted with a phosphite reagent to provide a site for the attachment of the cytidine analog to the oligonucleotide.

Referring to Fig. 3, the cytidine analog is then reacted with the oligonucleotide. Referring to Fig. 4, the protective groups $COCF_3$ and DMT are then removed and (Fig. 5) biotin added at the reactive NH_2 via an acylation reaction.

In more detail, the illustrated synthesis was carried out according to the following steps.

^{4}N-(3-aminopropyl)-2'-deoxycytidine (compound (3) of Fig. 2) was first prepared, as follows. A solution of sodium bisulfite (2M) and 1,3-diaminopropane (1M) was prepared by the addition of 1,3-diaminopropane (18.38g, 0.25M, 20.8ml) to an ice cold solution of sodium bisulfite (52.3g, 0.5M) in distilled water (200ml), the pH adjusted to 7 with concentrated HCl, and the volume made up to 250ml with distilled water.

2'-deoxycytidine monohydrate (5.0g, 20mM) was added and the reaction mixture warmed to 70°C and incubated overnight under argon. TLC in isopropan 1: ammonia:water (60:15:5) gave two spots of equal intensity (Rf = 0.29, 0.19). The reaction mixture was diluted 10-fold with 0.1M TRIS-HCl (30.28g, 0.25m, ph9) and the solution heated to 37°C. TLC as developed above gave one spot at Rf = 0.29 after one hour. The solution was concentrated to approximately 150ml and applied to a reversed phase C-18 column for desalting. The column was washed with two liters of distilled water and the product eluted with 10% acetonitrile in water to give compound (3) (2.0g).

Next, compound (3) was suspended in a mixture of dichloromethane (50ml)/triethylamine (4ml) and cooled to 4°C under argon. Trifluoroacetic anhydride (5.4g, 25mmol) was added dropwise by syringe over a ten minute period. The reaction mixture was stirred at room temperature for one hour and then an excess of trifluoroacetic anhydride (5.4g, 25mmol) was added. After one hour methanol (30ml) was added and the solvent was then removed <u>invacuo</u> yielding a yellow oil (approximately 15g), which was purified by column chromatography on silica gel (Baker Flash grade). Elution with dichloromethane:methanol (8:2) and evaporation of the solvent <u>invacuo</u> gave ^{4}N-(3-trifluoroacetylaminopropyl)-2'-deoxycytidine (compound (4) of Fig. 2).

Compound (4) (1g, 2.6mmol) was co-evaporated with anhydrous pyridine (3 x 25ml) and dissolved in anhydrous pyridine (30ml). 4,4'-dimethoxytritylchloride (0.97g, 2.8mmol) was added and the mixture stirred at room temperature for one hour. Methanol (30ml) was

added and then the reaction mixture was concentrated invacuo. The residue was partitioned between dichl romethane (100ml) and cold saturated sodium bicarbonate solution (100ml). The aqueous layer was further extracted with dichloromethane (100ml). The organic extracts were combined and dried using Na_2SO_4. Evaporation of the solvent invacuo gave a yellow oil which was purified by column chromatography on silica gel (Baker Flash Grade). Elution with dichloromethane/methanol/pyridine (950:49:1) and evaporation of the solvent invacuo gave 5'-O-(4,4-dimethoxytrityl)-^{4}N-(3-trifluoro-acetylaminopropyl)-2'-deoxycytidine (compound (5) of Fig. 2).

Compound (5) (1.3g, 1.9mmol) was dissolved in a mixture of chloroform (10ml, filtered through basic alumina) and diisopropylethylamine (1.3ml, 7.6mmol) under argon. To the ice cooled reaction mixture was added chloro-morpholino-methoxy-phosphine (0.29ml, 2.85mmol), dropwise, by syringe over three minutes. After 15 minutes at room temperature, the solution was transferred with 50ml ethylacetate into a 250ml separating funnel and extracted with cold, saturated sodium chloride solution (8 x 80ml). The organic layer was dried (Na_2SO_4) and then concentrated invacuo to yield a buff colored foam. The foam was dissolved in toluene (15ml) and added dropwise to hexane (200ml) at -80°C. The colorless precipitate, collected at room temperature and dried under a stream of argon, was 5'-O-(4,4-dimethoxytrityl)-3'-methoxymorpholinophosphine-^{4}N-(3-trifluoro-acetylaminopropyl)-2'-deoxycytidine (compound (6) of Fig. 2).

The next step was the incorporation of compound (6) int synthetic DNA. Compound (6), dissolved in acetonitrile (80mg ml^{-1}, 9.7 $umolsml^{-1}$) was used t functionalize oligonucleotides synthesised on an Applied Biosystems 380A DNA synthesiser, according to standard techniques. This resulted in functionalized oligonucleotide (7) of Fig. 3.

Following deprotection by treatment with thiophenol, ammonia and acetic acid to form the deprotected oligonucleotide (8) of Fig. 4, labeling with biotin was carried out as follows.

To oligonucleotide (8) (25 ug, 2.5 x 10^{-9}moles) dissolved in 0.1M sodium borate, pH 8.0, was added the N-hydroxysuccinimide ester of biotin amino caproic acid (23 m g, 5 x 10^{-8} moles) (Enzo Biochem) in dimethyl formamide (3 m l), and the reaction allowed to proceed for 4 hours. Residual biotin acylating reagent was removed from the biotinylated oligomer by chromotography on Sephadex G50. Further purification can be accomplished by reverse phase high performance liquid chromatography.

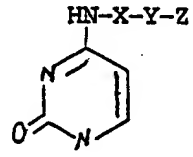
Use

The oligonucleotides which have been labeled using the cytidine analogs of the invention are used in nucleic acid hybridization reactions to detect nucleic acid sequences characteristic of specific bacteria, viruses, or cells, according to conventional techniques, e.g., as described in Landes U.S. Pat. Appln. S.N. 529,044, filed September 2, 1983, assigned to the same assignee as the present application, hereby incorporated by reference.

CLAIMS

1. An olig nucleotide containing at least ne cytidine analog labeled at C4 with a reporter group.

2. A cytidine analog in which the cytosine base portion thereof has the formula



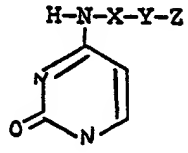
wherein Y is derived from a group capable of reacting with the N-hydroxy-succinimide ester of biotin amino caproic acid, Z is a protective group capable of preventing Y from reacting with a phosphite, and X is a linking group connecting Y and N.

3. The cytidine analog of claim 2 wherein X is $(CH_2)_n$, where n is an integer between 2 and 6, inclusive; and Z is $\overset{O}{\overset{\|}{C}}-F_3$.

4. The cytidine analog of claim 3, wherein n is 3.

5. A method of end-labelling an oligonucleotide chain, c mprising the steps of:

attaching a cytidine analog to said oligonucleotide chain, in which the cytosine base portion of said analog has the formula

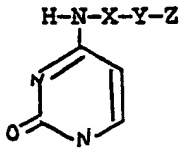


wherein Y is derived from a group capable of reacting with the N-hydroxy-succinimide ester of biotin amino caproic acid; Z is a protective group capable of preventing Y from reacting with phosphite, and X is a linking group connecting Y and N;

deprotecting said Y group; and

reacting said Y group with a reporter group.

6. A method of synthesizing a cytidine analog, in which the cytosine base portion of said analog has the formula:



wherein Y is derived from a group capable of reacting with the N-hydroxy-succinimide ester of biotin amino caproic acid; Z is a protective group capable of preventing Y from reacting with phosphite, and X is a linking group connecting Y and N;

comprising the steps of:

chemically adding said linking group to cytosine,

chemically adding said reactive group to said linking group; and

chemically adding said protective group to said reactive group.

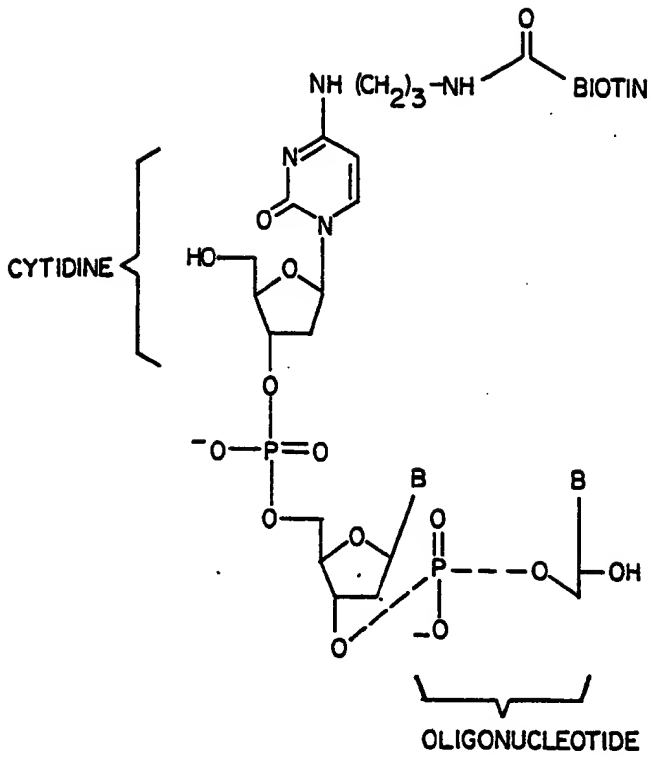


FIG 1

SUBSTITUTE SHEET

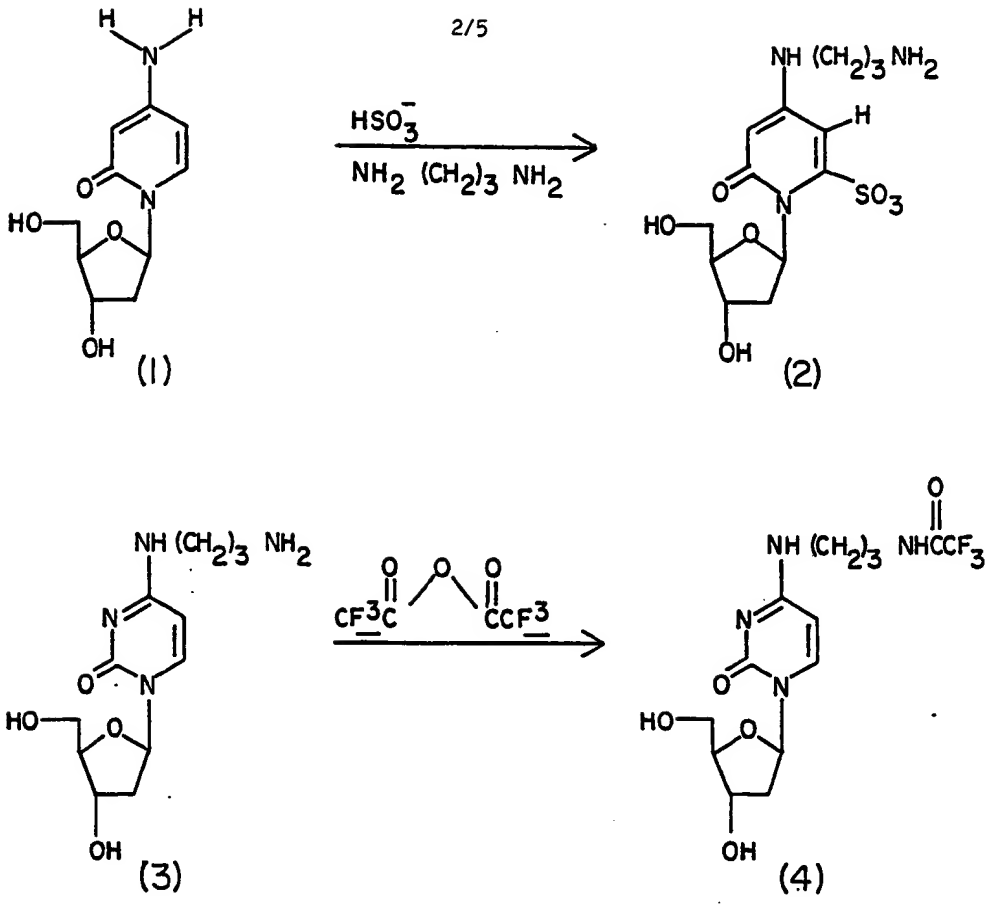


FIG 2

SUBSTITUTE SHEET

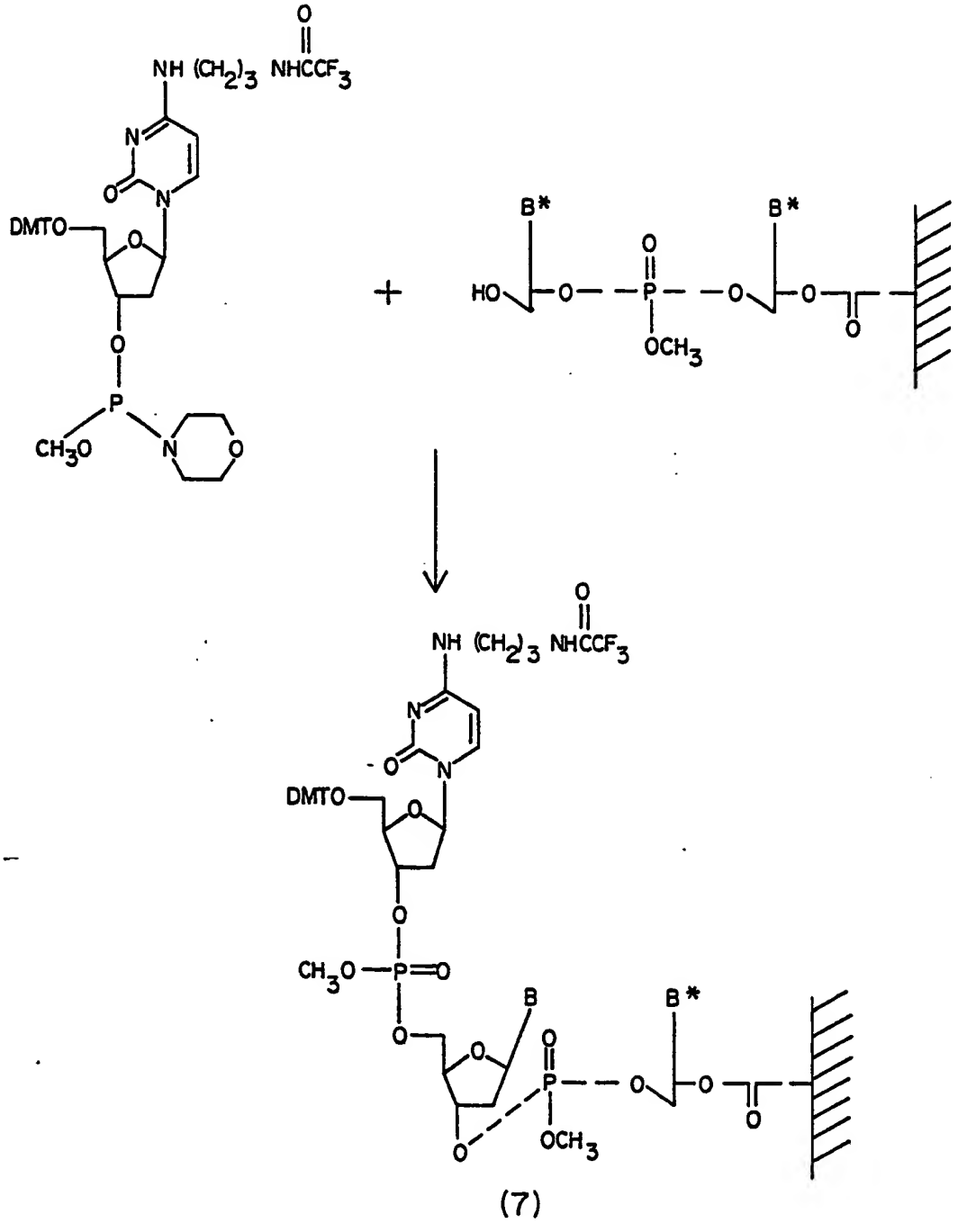


FIG 3

SUBSTITUTE SHEET

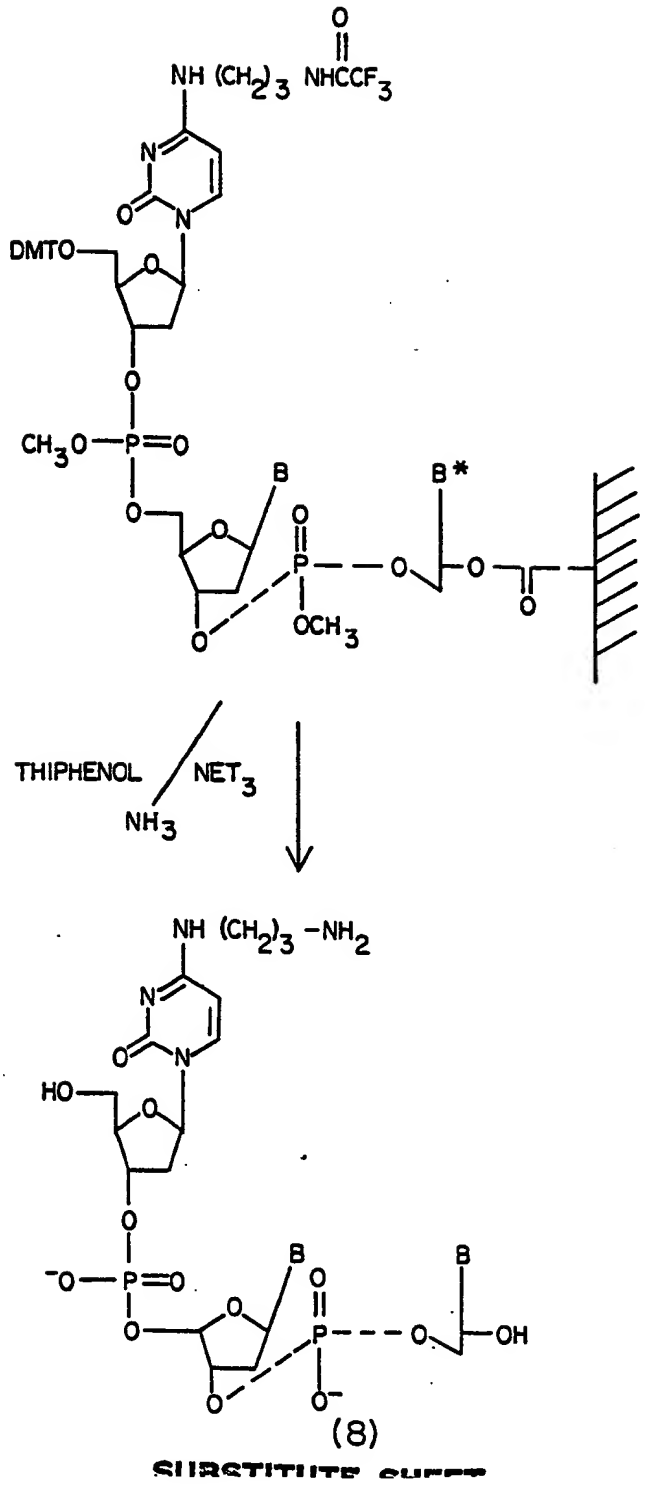

NH (CH2)3 NHCCF3
O
DMTO
CH3O—P=O
B
B*
OCH3
THIPHENOL
NET3
NH3
NH (CH2)3 –NH2
HO
–O—P=O
OH
(8)

FIG 4

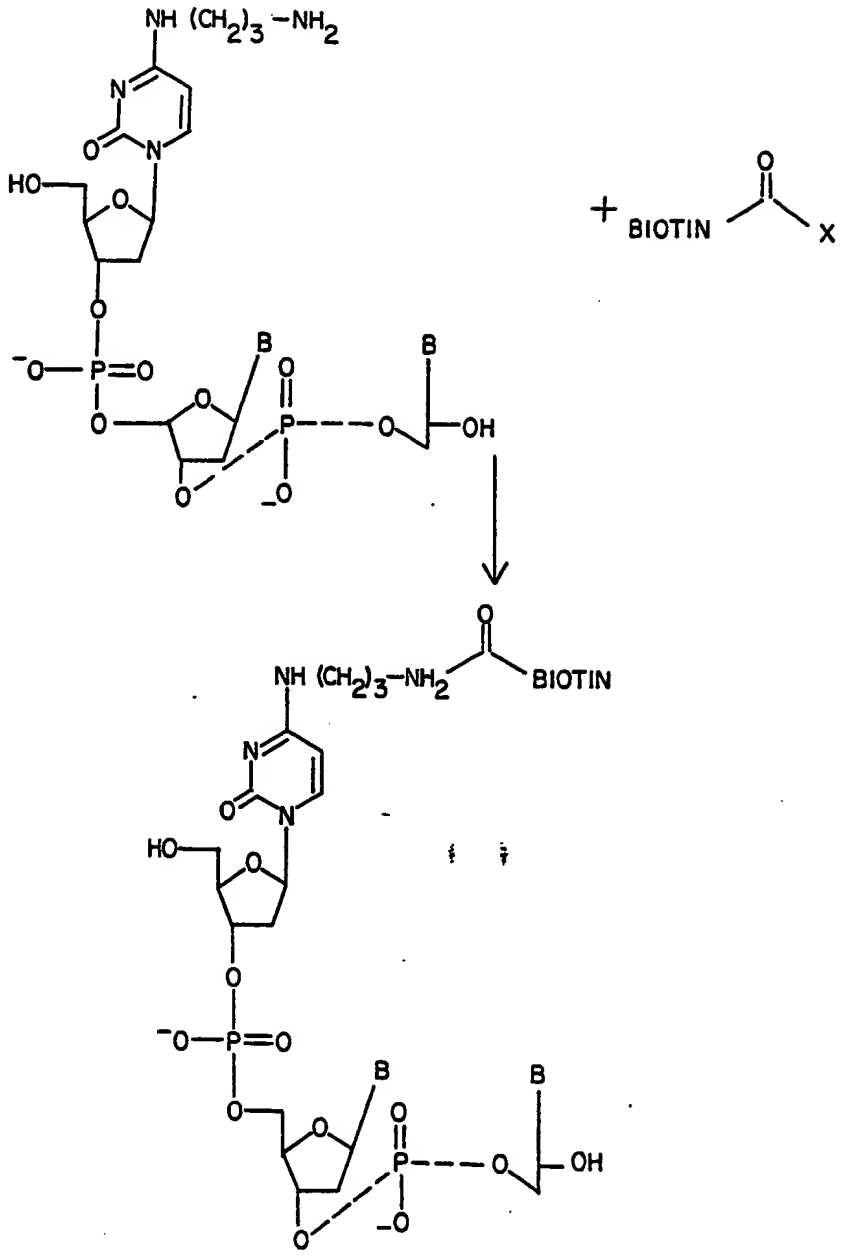


FIG 5

SUBSTITUTE SHEET

INTERNATI NAL SE[illegible]CH REPORT

International Application No PCT/US86/01052

I. CLASSIFICATION OF SUBJECT MATTER (if several classification symbols apply, indicate all) [3]

According to International Patent Classification (IPC) or to both National Classification and IPC

INT. Cl C07H 21/02 4th Edition
US Cl 536/22; 536/23; 536/29

II. FIELDS SEARCHED

Minimum Documentation Searched [4]

Classification System	Classification Symbols
US	536/22; 536/23; 536/29

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [5]

COMPUTER SEARCH: 4-Amine Substituted Cytidine
4-Amine Substituted Cytidine oligonucleotide

III. DOCUMENTS CONSIDERED TO BE RELEVANT [14]

Category *	Citation of Document, [16] with indication, where appropriate, of the relevant passages [17]	Relevant to Claim No. [18]
X	US, A, 3,847,898 (KELLY) published 12 November 1974, (columns 3-4)	2
A	US, A, 3,901,876 (VORBRUGGEN) published 26 August 1975	6
A	US, A, 4,097,665 (ISHIDA) published 27 June 1978	2-4
A	US, A, 4,310,662 (CREA) published 12, January 1982	1 and 5
X	US, A, 4,415,732, (CARUTHERS) published 15, November 1983 (columns 3-6)	1

* Special categories of cited documents: [15]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search [2]	Date of Mailing of this International Search Report [2]
16 JUNE 1986	30 JUL 1986
International Searching Authority [1]	**Signature of Authorized Officer** [20]
ISA/US	Johnnie R. Brown Johnnie R. Brown

Form PCT/ISA/210 (second sheet) (May 1986)